

LA COSTRUZIONE DI UN INDICATORE SINTETICO DI DOTAZIONE
INFRASTRUTTURALE: METODI E APPLICAZIONI A CONFRONTO

Claudio MAZZIOTTA¹, Matteo MAZZIOTTA², Adriano PARETO² e Francesco VIDOLI³

¹ Università degli Studi Roma Tre, via G. Chiabrera 190, 00152, Roma, c.mazziotta@uniroma3.it

² Istat, via A. Ravà 150, 00142, Roma, mazziott@istat.it; pareto@istat.it

³ Società per gli Studi di Settore S.p.A., via M. Maggini 47, 00147, Roma, fvidoli@sose.it

SOMMARIO

Il tema della sintesi di diversi indicatori in un indicatore rappresentativo dell'insieme dei fenomeni considerati è un tema ben noto in letteratura, soprattutto nel campo delle statistiche sociali. Spesso, tuttavia, la sintesi viene effettuata senza l'introduzione di alcuna ponderazione, dunque assumendo implicitamente che i diversi fenomeni considerati abbiano tutti lo stesso rilievo e possano quindi "entrare" nel procedimento di sintesi in termini paritetici. Questa assunzione, pur comprensibile per la difficoltà di evitare il rischio dell'arbitrarietà, potrebbe tuttavia condurre a risultati distorti, o quanto meno non completamente affidabili, dato che nella realtà difficilmente i fenomeni da sintetizzare possono essere posti su un livello paritetico.

La questione, peraltro, si complica quando si è in presenza di una doppia dimensione da sintetizzare: quella settoriale (nel lavoro qui presentato, le diverse categorie infrastrutturali) e quella territoriale (qui, le province italiane). Al problema – in particolare, a quello della costruzione di un indicatore di sintesi della dotazione infrastrutturale a livello di diverse unità territoriali – sono state date soluzioni diverse, tutte più o meno basate sulla variabilità, settoriale e/o territoriale, quale elemento da assumere per la determinazione dei pesi da introdurre eventualmente nell'algoritmo di sintesi.

Il lavoro che qui si presenta si pone l'obiettivo di mettere a confronto alcuni degli approcci a tal fine utilizzati e di verificare sia l'eventuale convergenza sia il grado di robustezza (e dunque di affidabilità) dei risultati ottenuti.

1 INTRODUZIONE: FINALITÀ E OGGETTO DEL LAVORO

Il procedimento di aggregazione di indicatori elementari in un indicatore di sintesi fornisce risultati diversi, come è ovvio, a seconda non soltanto dei metodi di aggregazione utilizzati, ma anche – e forse soprattutto – a seconda che sia inserito nella procedura di costruzione dell'indice sintetico un sistema di pesi. L'argomento è stato studiato sotto diverse angolazioni in letteratura: basti citare, tra i contributi più recenti, quello di Aiello e Attanasio (2004) e quello di Nardo *et al.* (2005), nel quale ultimo per conto dell'OCDE è stato predisposto una sorta di *manuale* sulla costruzione degli indici sintetici.

Il problema, per la verità, si complica quando alla dimensione settoriale degli indicatori elementari si affianca la dimensione territoriale, ossia quando agli indicatori sintetici viene richiesto di essere rappresentativi non soltanto di una, ma di entrambe le dimensioni. È questo il tipico problema che si incontra, ad esempio, nella costruzione di indicatori di sintesi della dotazione infrastrutturale, indicatori che si presentano articolati sia per categoria (trasporti, energia, etc.), sia per area (nel caso italiano, ad esempio, regioni o province). La maggior parte dei lavori che si sono occupati di questo problema avanzano soluzioni che, pur nella diversità degli approcci e degli algoritmi utilizzati, fanno in genere riferimento alla variabilità – territoriale o settoriale o entrambe congiuntamente – quale caratteristica da assumere alla base della determinazione dei pesi da introdurre nella procedura di sintesi (si vedano, ad esempio, i lavori di Brunini *et al.*, 2004; Mazziotta C. e Vidoli, 2007; Mazziotta M. e Pareto, 2007).

Il lavoro qui presentato si inserisce in questo contesto metodologico. Più specificamente, esso intende fare il punto su alcuni approcci di ponderazione utilizzati per la costruzione degli indicatori sintetici riferiti alla dotazione infrastrutturale, approcci sperimentati, prevalentemente anche se non esclusivamente, dagli autori del presente *paper*. Inoltre, si intende condurre un riscontro empirico sui risultati ottenibili dai diversi approcci sfruttando l'opportunità dei dati statistici di recente forniti dall'ISTAT sulle singole categorie infrastrutturali a livello delle 103 province italiane (ISTAT, 2006).

Assumendo come approccio-base il metodo della sintesi non ponderata attraverso la media geometrica degli indicatori elementari (in sigla, GEOM), sono stati posti a confronto i tre metodi di sintesi ponderata di seguito indicati, i quali, come già accennato, fondano la determinazione del sistema di pesi sulla variabilità dei dati elementari, sia pure variamente intesa e quantificata:

- *metodo tassonomico di Wroclaw (TAX)*;
- *metodo denominato delle penalità per coefficiente di variazione (MPCV)*;

– metodo noto come *Benefit of the Doubt (BOD)*.

Vale la pena di sottolineare che obiettivo del lavoro *non* è di stabilire quale approccio sia *migliore* o comunque *preferibile* ad un altro. Così formulato, si tratta infatti di un problema mal posto, dato che approcci diversi possono giustificarsi in ragione di obiettivi diversi posti a base dell'analisi: l'ottica del riequilibrio regionale, ad esempio, è diversa da quella della più efficiente allocazione delle risorse; pertanto, a seconda che si privilegi l'una o l'altra impostazione, è più corretto utilizzare un approccio di sintesi piuttosto che un altro.

L'obiettivo che ci si pone è più specifico e circoscritto, ma forse, proprio per questo, di maggiore interesse per gli utilizzatori degli indicatori in questione: esso mira a verificare la robustezza dei risultati ottenuti nei diversi casi in cui si voglia utilizzare un approccio di aggregazione basato su una procedura di sintesi ponderata. Il lavoro, in altri termini, attraverso la misura delle differenze riscontrate tra i risultati ottenuti con approcci diversi di aggregazione, intende saggiare la sensibilità dei risultati stessi e, in definitiva, l'affidabilità "statistica" dei metodi utilizzati, al di là della loro appartenenza ad ottiche economiche diverse.

Il lavoro è articolato nel modo seguente: nel paragrafo 2 vengono brevemente descritti i metodi di sintesi utilizzati; nel paragrafo 3 si procede ad effettuare il confronto tra i risultati ottenuti, dapprima in termini descrittivi, poi con l'ausilio di analisi della "sensitività" e di robustezza; il paragrafo 4, infine, sintetizza le conclusioni cui si perviene al termine del lavoro.

2 I DIVERSI METODI DI SINTESI PRESI IN CONSIDERAZIONE

2.1 Approccio base: sintesi attraverso media geometrica non ponderata

La metodologia più frequentemente¹ seguita per costruire indicatori sintetici di dotazione infrastrutturale sulla base di dati fisici (ossia non valutati in termini monetari) può sintetizzarsi nelle seguenti fasi: *i)* raccolta dei dati disponibili e costruzione degli indicatori elementari; *ii)* normalizzazione degli indicatori elementari; *iii)* standardizzazione degli indici normalizzati; *iv)* aggregazione degli indicatori standardizzati in modo da passare dalle categorie infrastrutturali elementari alle categorie principali ed infine ad un indicatore di sintesi dell'intera dotazione.

¹ Tale metodologia è stata inizialmente messa a punto in una ricerca sulla dotazione infrastrutturale promossa dalla Commissione Europea (Biehl *ed.*, 1986) ed è stata successivamente confermata, sia pure con alcuni adattamenti e aggiustamenti, in diversi altri studi realizzati nei successivi venti anni: tra i più recenti, si vedano Di Palma e Mazziotta (2003) e Mazziotta (2006).

I passi costitutivi di tale procedura possono sintetizzarsi nei termini seguenti:

- i) *costruzione degli indicatori elementari*. La disponibilità statistica di base condiziona ovviamente sia il dettaglio delle categorie infrastrutturali da quantificare, sia anche il livello delle aree territoriali da prendere in considerazione;
- ii) *normalizzazione degli indicatori elementari*. Poiché i dati rilevati nella fase precedente, espressi nelle più diverse unità di misura, non consentono alcuna comparazione tra le diverse aree considerate, si deve procedere alla normalizzazione² degli indicatori elementari, operazione che in genere viene effettuata rapportando gli indicatori elementari ad un numerario costituito dalla superficie territoriale dell'area interessata (infrastrutture "a rete" o *space serving*) o dalla popolazione (indicatori "puntuali" o *population serving*);
- iii) *standardizzazione degli indici normalizzati*. Gli indicatori così ottenuti non sono ancora comparabili, poiché risultano espressi in unità di misura diverse (km di strade per kmq di superficie e numero di telefoni per abitante, ad esempio). La standardizzazione è appunto finalizzata ad ottenere indicatori che siano depurati dalle specifiche unità di misura: uno dei metodi più utilizzati a tal fine è quello di rapportare, all'interno della stessa categoria infrastrutturale, gli indici normalizzati di ciascuna unità territoriale al valore massimo della serie, ottenendo in tal modo scale della medesima ampiezza (tra 0 e 1, per l'appunto, o tra 0 e 100, come più comunemente viene fatto) e quindi l'immediata comparabilità tra gli indicatori considerati;
- iv) *aggregazione degli indicatori standardizzati*. Il procedimento più consueto consiste nel ricorso ad una qualche forma di media, spesso identificata nella media aritmetica per l'aggregazione degli indicatori elementari all'interno di una categoria principale e nella media geometrica per la sintesi tra le categorie principali. Il ricorso alle due diverse medie assume per implicito che vi sia una certa fungibilità tra le categorie elementari appartenenti alla stessa categoria ed invece una minore sostituibilità tra le categorie principali che concorrono a determinare l'indice sintetico complessivo.

Il procedimento sopra richiamato – nella versione della media geometrica – viene qui assunto come *approccio-base*: ciò senza nessun particolare giudizio di valore, ma semplicemente per avere un termine di riferimento rispetto al quale verificare gli scostamenti nei risultati ottenuti con gli approcci di sintesi ponderata di seguito illustrati.

² E' appena il caso di rilevare che i termini di *normalizzazione* e *standardizzazione* sono, come di consueto, convenzionali: nell'accezione qui indicata essi sono stati utilizzati negli studi citati nella nota precedente. In altri contesti, tuttavia, la terminologia è impiegata con significato del tutto "rovesciato": si parla cioè di standardizzazione per l'operazione qui definita come normalizzazione e viceversa.

2.2 Sintesi con il metodo tassonomico di Wroclaw

Si tratta di una tecnica elaborata originariamente presso l'Università di Wroclaw, che ha conosciuto una discreta diffusione anche nel nostro paese, soprattutto per l'elaborazione di indicatori economico-sociali (si vedano, ad esempio, Schifini D'Andrea, 1982; Quirino, 1990; Mazziotta, 1998). Essa si basa su un principio molto semplice, che nel caso in esame potrebbe esprimersi così: l'area migliore è quella che presenta la minor distanza da un'area "ideale", contraddistinta dal fatto di avere le migliori *performance* per tutte le variabili considerate (nel nostro caso, di presentare il livello più elevato per ciascuna delle categorie infrastrutturali considerate). Sulla base del calcolo delle distanze (euclidee) di tutte le province dalla provincia "ideale", è possibile costruire una graduatoria in cui le diverse province risultano ordinate rispetto alla maggiore o minore distanza dalla situazione ottimale, assunta come misura rispettivamente del maggiore o minore deficit di dotazione infrastrutturale.

Formalmente, si parte dalla matrice originaria dei dati $X=\{x_{ij}\}$, la si trasforma nella corrispondente matrice standardizzata $Z=\{z_{ij}\}$, il cui generico elemento è dato da $z_{ij} = (x_{ij} - \bar{x}_j) / \sigma_j$ e si procede all'individuazione del cosiddetto "modello estremo", costituito dal vettore riga contenente i valori massimi degli indicatori relativi a tutte le categorie infrastrutturali considerate. Tale vettore identifica la provincia "ideale", così definita perché caratterizzata dalla migliore dotazione riscontrata per ogni categoria di infrastrutture esaminata.

Rispetto a tale vettore, che definisce dunque la provincia "ideale" per livello di dotazione infrastrutturale:

$$z_o = (z_{o1}, \dots, z_{oj}, \dots, z_{om})$$

vengono successivamente calcolate, ad esempio con metrica di tipo euclideo, le distanze di ciascuna provincia:

$$D_{oi} = \sqrt{\sum_{j=1}^m (z_{ij} - z_{oj})^2}.$$

Naturalmente, più alto sarà il livello di D_{oi} e maggiore sarà per la provincia i il divario infrastrutturale. Poiché D_{oi} è ancora una misura assoluta di divario, si rende opportuno trasformarla in una misura relativa sulla base dell'espressione $d_i = D_{oi} / D_o$, in cui $D_o = \bar{D}_o + 2\sigma_o$, laddove \bar{D}_o è la media di tutte le distanze e σ_o è lo scostamento quadratico medio delle distanze stesse. L'indice d_i costituisce dunque il parametro ordinatore della graduatoria di sintesi delle province rispetto al diverso livello di dotazione infrastrutturale da esse presentato: tale indice assume valore nullo quando la distanza tra la provincia i e il

modello estremante è nulla, ossia quando tale provincia coincide con la provincia “ideale”; non è invece definito il limite superiore che comunque, data la struttura dell’indice, è prossimo all’unità.

Con questo metodo di sintesi si attua implicitamente una ponderazione delle variabili elementari, che risultano tanto più influenti sull’indice sintetico quanto più notevoli sono le distanze registrate rispetto alla situazione “ideale”.

2.3 Sintesi con il metodo delle penalità per coefficiente di variazione

Il metodo delle penalità per coefficiente di variazione (Mazziotta M. e Pareto, 2007) si propone di fornire una misura sintetica della dotazione infrastrutturale di un insieme di unità territoriali nell’ipotesi che ciascuna componente della dotazione non sia sostituibile con le altre o lo sia solo in parte. Ciò comporta l’introduzione di una “penalità” per le unità che presentano una dotazione non bilanciata di tutte le componenti.

Il metodo per il calcolo dell’indice sintetico prevede i seguenti passi.

a) Standardizzazione degli indicatori

Data una matrice $X=\{x_{ij}\}$ di n righe (unità territoriali) e m colonne (indicatori), si passa alla matrice $Z=\{z_{ij}\}$ in cui:

$$z_{ij} = \frac{(x_{ij} - M_{x_j})}{S_{x_j}} S + M$$

dove:

$$M_{x_j} = \frac{\sum_{i=1}^n x_{ij}}{n}; \quad S_{x_j} = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - M_{x_j})^2}{n}};$$

x_{ij} = valore dell’indicatore j nell’unità i;

$S = 10$ e $M = 100$.

b) Calcolo della “variabilità orizzontale”

Data la matrice $Z=\{z_{ij}\}$, si calcola il vettore dei coefficienti di variazione $CV=\{cv_i\}$ in cui:

$$cv_i = \frac{S_{z_i}}{M_{z_i}}$$

dove:

$$M_{z_i} = \frac{\sum_{j=1}^m z_{ij}}{m} \quad \text{e} \quad S_{z_i} = \sqrt{\frac{\sum_{j=1}^m (z_{ij} - M_{z_i})^2}{m}}.$$

c) Costruzione dell'indice sintetico

L'indice sintetico dell'i-esima unità (MPCV_i) si ottiene mediante la formula:

$$\text{MPCV}_i = M_{z_i} (1 - cv_i^2) = M_{z_i} - S_{z_i} cv_i,$$

in cui si corregge la media aritmetica degli indicatori standardizzati sottraendo una quantità proporzionale allo scostamento quadratico medio e funzione diretta del coefficiente di variazione.

L'indice ottenuto è tanto maggiore quanto più grande è la media aritmetica degli indicatori standardizzati e quanto più piccolo è lo scostamento quadratico medio.

Il metodo delle penalità per coefficiente di variazione si basa su un modello additivo e non richiede, come avviene per la media geometrica, che l'intensità totale del fenomeno (la dotazione infrastrutturale) sia uguale al prodotto delle singole componenti.

La procedura di standardizzazione consente di liberare gli indicatori elementari sia dall'unità di misura che dalla loro variabilità e non comporta la definizione di un vettore di valori obiettivo (unità "ideale"), in quanto sostituisce tale vettore con l'insieme dei valori medi³. In tal modo, risulta agevole individuare le unità territoriali che hanno un livello di dotazione delle infrastrutture al di sopra di quello medio (valori maggiori di 100) e le unità che hanno un livello di dotazione al di sotto della media (valori minori di 100).

La funzione di aggregazione (media aritmetica dei valori standardizzati) viene corretta mediante un coefficiente di penalità che dipende, per ciascuna unità territoriale, dalla variabilità degli indicatori rispetto al valor medio ("variabilità orizzontale"). Ciò consente di penalizzare il punteggio delle unità che, a parità di media aritmetica, hanno un maggiore squilibrio tra i valori degli indicatori.

L'indice MPCV è di semplice determinazione, non risente della presenza di dati anomali (*outliers*) ed è facilmente interpretabile e comparabile nel tempo.

³ Il criterio di determinazione dell'unità "ideale" è spesso arbitrario e soggettivo, non è univoco e può variare nel tempo. Un'ulteriore perplessità sorge in merito all'eventualità che l'unità "ideale" a cui si perviene possa, di fatto, non esistere, dal momento che gli indicatori impiegati sono, nella quasi totalità dei casi, tutt'altro che indipendenti e i valori "ottimali" potrebbero costituire una combinazione incongruente (Aureli Cutillo, 1996).

2.4 Sintesi con il metodo *Benefit of the Doubt*

La metodologia *Benefit of the Doubt* (BOD), originariamente proposta per valutare le performance macroeconomiche (Melyn and Moesen, 1991) e recentemente adattata alla teoria degli indicatori (cfr. Cherchye *et al.* 2004, 2005, 2006), è un'applicazione della *Data Envelopment Analysis* (DEA).

L'indicatore composito è, secondo questa logica, espresso come rapporto tra la performance della singola unità con quella del *benchmark*, facendo sì che il set ottimo di pesi (se esiste) garantisca ad ogni unità associata la migliore posizione possibile rispetto a tutte le altre.

I pesi ottimi sono ottenuti risolvendo il seguente problema:

$$CI_c^* = \max_{w_{cq}} \frac{\sum_{q=1}^Q I_{cq} w_{cq}}{\max_{k \in [1...C]} \sum_{q=1}^Q I_{kq} w_{kq}}, \forall c = 1...C$$

dove i pesi sono posti non negativi e I_{cq} rappresenta il punteggio normalizzato del q-esimo indicatore semplice ($q=1,...,Q$) per l'unità c ($c=1,...,C$) e w_{cq} il peso corrispondente. L'indicatore composito risultante varierà quindi tra zero (la performance più bassa) e 1 (il benchmark).

Utilizzando i metodi della ricerca operativa, l'equazione sopra riportata può essere ridotta ad un problema di programmazione lineare, aggiungendo il vincolo della somma ad 1, moltiplicando tutti i pesi per un fattore comune (il che non altera il valore dell'indicatore composito) e risolvendo tramite noti algoritmi di ottimizzazione lineare:

$$CI_c^* = \max_{w_{cq}} \sum_{q=1}^Q I_{cq} w_{cq}$$

sotto i vincoli:

$$\begin{aligned} \sum_{q=1}^Q I_{cq} w_{cq} &\leq 1 \\ w_{cq} &\geq 0 \\ \forall c &= 1...C; \forall q = 1...Q \end{aligned}$$

Questo approccio offre molteplici vantaggi tra i quali quello che i pesi sono determinati endogenamente dai dati osservati e che, di conseguenza, il *benchmark* non è scelto in base a

modelli teorici, ma è semplicemente una combinazione lineare delle migliori *performance* osservate (sotto questo profilo si nota una certa somiglianza dell'approccio BOD con le ipotesi a base del metodo tassonomico).

Il limite di questo approccio è legato direttamente alla soluzione del problema DEA: dato che i pesi sono specifici per ogni unità, confronti tra unità non sono possibili ed il valore specifico dell'indicatore composito dipende dalla performance del *benchmark*.

Il problema, nella sua formulazione iniziale, nasconde però un rischio di molteplicità delle soluzioni, con la conseguenza che i pesi non risulterebbero univocamente determinati; si noti inoltre che il problema delle soluzioni multiple dipende dal *set* di vincoli imposti al problema di massimizzazione: più ampio è il *range* di variazione dei pesi e più bassa sarà la probabilità di ottenere una soluzione univoca.

Usualmente si utilizzano informazioni derivate da opinioni di esperti per cercare di inserire vincoli appropriati sui pesi. Nell'approccio utilizzato in una recente applicazione alla dotazione infrastrutturale (Mazziotta C. e Vidoli, 2007), i vincoli sui pesi sono, al contrario, determinati endogenamente, prendendo come idea base quella di pesare ogni singolo indicatore in base ai rapporti tra il limite inferiore e superiore degli intervalli di confidenza delle varianze campionarie degli indicatori semplici.

Partendo dall'ipotesi, quindi, che ogni indicatore sia descritto da una curva gaussiana, si può calcolare l'intervallo di confidenza della varianza campionaria:

$$P\left(\frac{C-1}{\chi^2_{C-1,1-\alpha/2}} S^2 < \sigma^2 < \frac{C-1}{\chi^2_{C-1,\alpha/2}} S^2\right) = 1 - \alpha$$

$$\text{Prob}\left\{low_{I_q} \leq \bar{S}^2 \leq high_{I_q}\right\} = 1 - \alpha$$

Possiamo quindi utilizzare low_{I_q} e $high_{I_q}$ per ogni indicatore per ricostruire i tassi marginali di sostituzione tra indicatori nella forma:

$$\frac{low_{I_i}}{high_{I_j}} \leq \frac{w_{I_i}}{w_{I_j}} \leq \frac{high_{I_j}}{low_{I_i}}, \forall i, j = 1 \dots Q$$

Applicando questi vincoli sui pesi nell'equazione iniziale, si ottiene l'indicatore di sintesi ricercato.

3 CONFRONTO TRA I RISULTATI OTTENUTI CON I DIVERSI METODI DI SINTESI PONDERATA

3.1 Premessa

In questo paragrafo sono illustrati i risultati ottenuti applicando i diversi metodi di sintesi sopra illustrati ad un insieme di indicatori di dotazione infrastrutturale tratto da una recente pubblicazione della statistica ufficiale (Istat, 2006). Per brevità – ed anche per il carattere essenzialmente metodologico del presente lavoro – l’applicazione è stata condotta su un *set* ristretto di categorie infrastrutturali, concernente i trasporti terrestri, articolati in trasporti stradali e trasporti ferroviari. Gli indicatori utilizzati sono identificati nella Tabella 1 e riportati, nella loro specifica dimensione quantitativa, nel Prospetto 1 in Appendice.

Dai semplici indici di variabilità ivi riportati si rileva che i sei indicatori presentano tutti una variabilità piuttosto accentuata, evidenziata da un coefficiente di variazione che per la metà dei casi risulta compreso tra il 30 e il 40%, e per l’altra metà oscilla intorno al 100%, in due casi addirittura superandolo. Questa caratteristica dei dati di base si rivela molto opportuna ai fini dell’obiettivo che qui ci si propone: in presenza, infatti, di un’elevata variabilità di partenza la verifica di robustezza degli approcci di sintesi utilizzati affronterà una prova più impegnativa e l’eventuale suo esito positivo sarà più convincente, proprio perché meno facilitato dalla distribuzione dei dati di base.

Tabella 1 Indicatori di dotazione delle infrastrutture dei trasporti stradali e ferroviari

Denominazione	Codice Istat
Trasporti stradali	
Chilometri di strade comunali per 10 kmq di superficie territoriale	ITR012
Chilometri di strade provinciali per 100 kmq di superficie territoriale	ITR013
Chilometri di strade statali per 100 kmq di superficie territoriale	ITR014
Chilometri di autostrade per 1.000 kmq di superficie territoriale	ITR015
Trasporti ferroviari	
Chilometri di rete ferroviaria F.S. elettrificata a binario doppio per 1.000 kmq di superficie territoriale	ITR064
Chilometri di rete ferroviaria F.S. elettrificata a binario semplice per 1.000 kmq di superficie territoriale	ITR065

Fonte: ISTAT (2006)

3.2 Classifiche, graduatorie e cartografie

Il confronto tra quattro metodologie così diverse, al fine di comprendere le rispettive peculiarità e quindi le reciproche differenze, è stato impostato partendo da semplici analisi descrittive e grafiche. La Tabella 2 rappresenta una matrice in cui è stata calcolata la differenza media assoluta di rango e l'indice di Spearman tra le graduatorie delle province italiane che si ottengono dall'applicazione dei diversi approcci descritti nel precedente paragrafo 2, incluso quello della media geometrica non ponderata, assunto come termine di riferimento di base.

Tabella 2 Differenza media assoluta di rango e indice di Spearman tra le graduatorie delle province italiane ottenute dall'applicazione di diversi approcci di costruzione dell'indicatore di sintesi (infrastrutture dei trasporti stradali e ferroviari)

Metodo di sintesi	Media geometrica	Indice MPCV	Metodo di Wroclaw	Indice BOD
Media geometrica				
- Differenza media assoluta	0,000	0,835	2,757	4,893
- Indice di Spearman	1,000	0,999	0,991	0,976
Indice MPCV				
- Differenza media assoluta	-	0,000	2,660	4,796
- Indice di Spearman	-	1,000	0,992	0,976
Metodo di Wroclaw				
- Differenza media assoluta	-	-	0,000	6,350
- Indice di Spearman	-	-	1,000	0,954
Indice BOD				
- Differenza media assoluta	-	-	-	0,000
- Indice di Spearman	-	-	-	1,000

I valori della matrice offrono interessanti spunti di riflessione, in realtà più per quanto riguarda le differenze medie assolute che per gli indici di cograduazione; questi ultimi, infatti, sono tutti maggiori di 0,95, dimostrando una elevata concordanza tra le graduatorie ottenute con i quattro approcci. Si segnala, in particolare, l'uniformità di comportamento (indice di Spearman pari a 0,999) tra GEOM e MPCV: ciò è dovuto alla costruzione stessa di quest'ultimo metodo, in quanto esso usa una funzione di penalità "orizzontale" a correzione della media aritmetica.

L'analisi delle differenze medie assolute rivela di quante posizioni di graduatoria, in media, le province italiane mutano quando si passa da un metodo ad un altro. Anche in questo caso emerge la "vicinanza" tra MPCV e GEOM che, in media, differiscono di 0,835 posizioni. Il metodo TAX si distanzia dai sopracitati due metodi di circa 2,7 posizioni e dal BOD di 6,350 posizioni (massima distanza). Lo stesso indice BOD presenta una differenza media assoluta con MPCV e con GEOM di circa 4,8 posizioni.

I risultati ottenuti possono essere meglio apprezzati dall'evidenza visiva di un cartogramma, opportunamente articolato a livello di provincia (Figura 1). A tal fine, la distribuzione dei valori dell'indice sintetico ottenuta da ciascuno dei quattro approcci considerati è stata suddivisa in quintili, in modo da avere classi con la stessa frequenza, identificate con gli stessi caratteri cromatici. Ovviamente, l'esame dei cartogrammi conferma quanto già emerso sulla base degli indici riportati in Tabella 2, ossia una buona convergenza tra i risultati ottenuti con i diversi metodi. E infatti le province che cambiano di classe, nel passaggio da un metodo all'altro, sono davvero poche; una maggiore uniformità, comunque, si nota tra le rappresentazioni cartografiche risultanti dall'applicazione degli approcci GEOM, TAX e MPCV, mentre più distanziato dagli altri appare il risultato ottenuto con il metodo BOD.

3.3 Analisi della sensitività

In questo paragrafo sono riportati i risultati di un'analisi di "sensitività" degli indici di dotazione infrastrutturale ottenuti con i quattro metodi. In particolare, si è voluto studiare se e con quanta intensità cambiano le graduatorie delle province italiane a seguito dell'eliminazione di un indicatore elementare dall'insieme di partenza.

Dati i K indicatori elementari ($K=6$), sono state condotte K repliche eliminando, ogni volta, un indicatore diverso e calcolando i valori degli indici sintetici sulla base dei $K-1$ indicatori rimanenti. Per ogni replicazione, sono state costruite le graduatorie delle province italiane secondo i vari metodi e, per ciascuna provincia, sono state calcolate le differenze assolute di rango tra la posizione nella graduatoria originaria e la posizione nella graduatoria relativa ai $K-1$ indicatori.

Nella Tabella 3 sono riportate le principali caratteristiche, per i quattro metodi considerati, delle distribuzioni delle differenze assolute di rango delle province italiane, distribuzioni ottenute di volta in volta a seguito dell'eliminazione di un indicatore elementare dal *set* dei 6 originari.

Figura 1 Cartogrammi degli indicatori di sintesi ottenuti dall'applicazione dei quattro approcci utilizzati (province italiane)

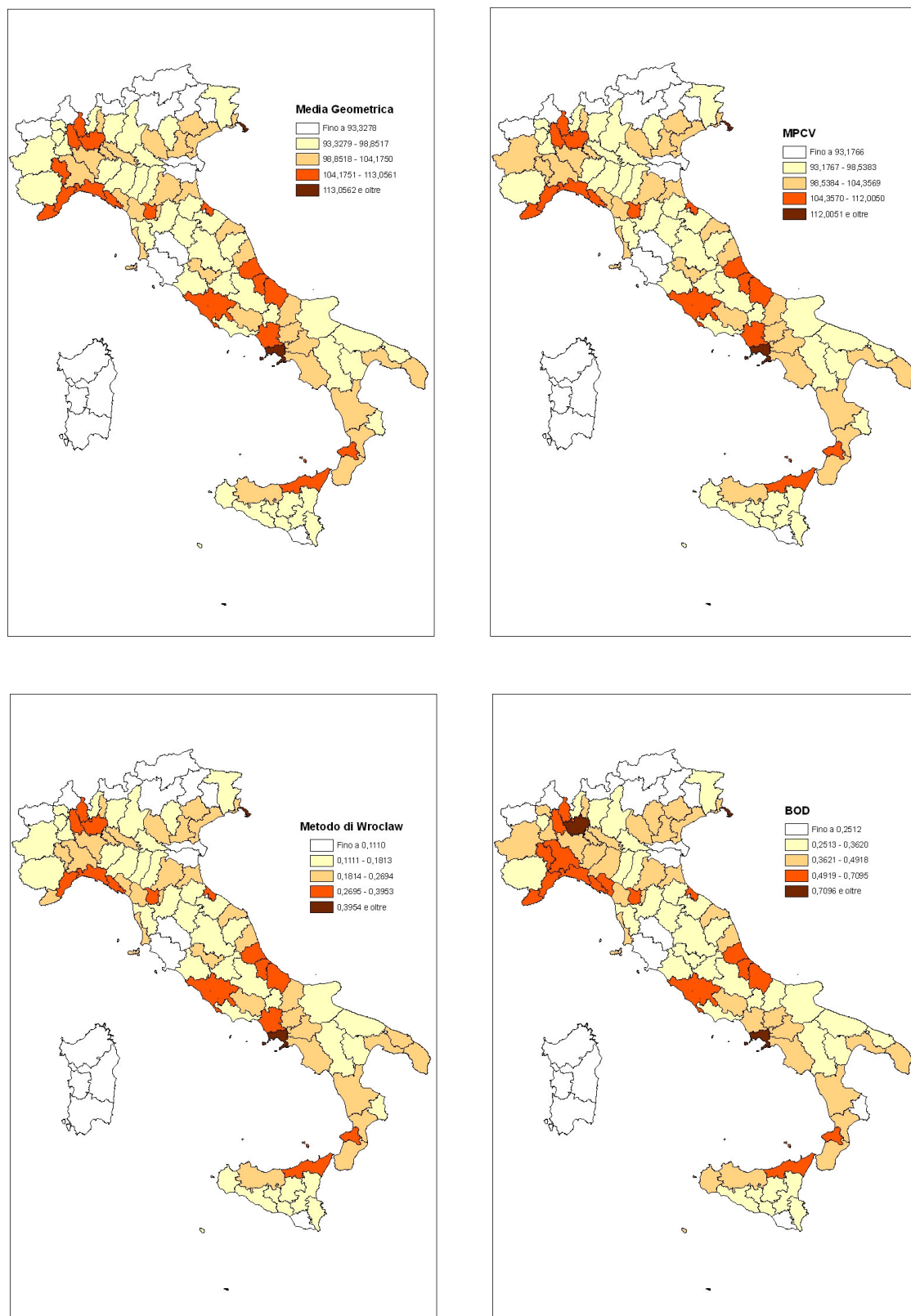


Tabella 3 Media, scostamento quadratico medio e campo di variazione degli spostamenti in graduatoria delle province italiane, a seguito dell'eliminazione di un indicatore per volta dal *set* originario, per i quattro metodi di sintesi

Indicatore	Media geometrica	Indice MPCV	Metodo di Wroclaw	Indice BOD
Media delle differenze assolute di rango				
ITR012	5,46	5,57	4,10	5,56
ITR013	5,50	5,51	5,71	4,64
ITR014	6,45	6,21	7,05	2,80
ITR015	5,83	6,08	4,29	9,73
ITR064	4,83	4,91	7,53	3,16
ITR065	6,04	5,94	6,80	5,35
Media	5,68	5,71	5,91	5,21
S.q.m.	0,51	0,43	1,33	2,27
S.q.m. delle differenze assolute di rango				
ITR012	5,77	5,71	4,07	5,38
ITR013	5,59	5,26	5,73	4,40
ITR014	5,41	5,39	6,52	2,78
ITR015	5,61	5,80	4,18	9,10
ITR064	4,65	4,70	6,54	3,12
ITR065	6,31	5,98	6,46	6,06
Media	5,56	5,47	5,58	5,14
S.q.m.	0,49	0,42	1,07	2,11
Campo di variazione delle differenze assolute di rango				
ITR012	37,00	37,00	21,00	41,00
ITR013	29,00	26,00	29,00	26,00
ITR014	27,00	25,00	36,00	15,00
ITR015	24,00	25,00	22,00	46,00
ITR064	20,00	18,00	26,00	16,00
ITR065	35,00	30,00	36,00	34,00
Media	28,67	26,83	28,33	29,67
S.q.m.	5,91	5,76	6,02	11,76

Dall'esame dei risultati risulta evidente la maggiore sensitività agli indicatori elementari dell'indice BOD che, pur attestandosi su una differenza media assoluta di rango pari a 5,21 posizioni, presenta tuttavia un'escursione piuttosto ampia, da un valore minimo di 2,8 posizioni, corrispondente all'eliminazione dell'indicatore ITR014, ad un massimo di 9,73 posizioni, corrispondente all'eliminazione dell'indicatore ITR015.

Per contro, la media geometrica (GEOM) e il metodo delle penalità per coefficiente di variazione (MPCV) si attestano su valori medi leggermente superiori a quelli del BOD (5,68 e 5,71), ma fanno registrare una maggiore stabilità dei risultati al variare degli indicatori. Ciò è dovuto, principalmente, alla procedura di standardizzazione degli indicatori che consente di

attribuire lo stesso peso alle diverse variabili, avendole depurate dalla variabilità. Il metodo tassonomico (TAX), infine, si pone in una posizione intermedia, con uno spostamento medio in graduatoria di 5,91 posizioni.

L'analisi degli scostamenti quadratici medi e dei campi di variazione delle differenze assolute di rango conferma la maggiore variabilità degli spostamenti in graduatoria che si verificano con il metodo BOD: infatti, eliminando l'indicatore ITR015, lo scostamento quadratico medio ammonta a 9,10 posizioni, contro le 4,18 del metodo tassonomico, e il campo di variazione degli spostamenti in graduatoria è pari a 46 posizioni.

Nel complesso, la media geometrica e l'indice MPCV forniscono risultati molto simili, anche se l'uso della "penalità", abbinato alla media aritmetica, sembrerebbe ridurre la sensitività in termini di variabilità delle distribuzioni.

3.4 Analisi di robustezza

Nel presente paragrafo viene svolta un'analisi di robustezza sulle stime dell'indicatore di sintesi ottenuto dalle quattro metodologie considerate, puntando l'attenzione non tanto, come nel paragrafo precedente, sulla corrispondenza tra indicatore di sintesi e fenomeno da studiare (nel nostro caso la dotazione infrastrutturale nel campo dei trasporti), quanto piuttosto sulla più o meno ampia modificazione dei risultati al variare di un disturbo casuale aggiunto. In tal modo i metodi in questione vengono verificati sulla base della loro robustezza intrinseca, e dunque della loro più o meno ampia affidabilità "statistica".

Più precisamente, si è proceduto nel modo seguente:

- per ogni indicatore k , ai dati originari è stato sommato un disturbo casuale uniforme, espresso dalla formula:

$$I_k^N = I_k + Unif(-\alpha, \alpha) \cdot \overline{I_k}, \forall k = 1, \dots, K$$

con α inizialmente pari a 0,05 e $\overline{I_k}$ valor medio dell'indicatore k^4 ;

- per ogni estrazione casuale⁵, sono stati dapprima calcolati K indicatori semplici con valori nell'intorno degli originali, e successivamente dai quattro approcci qui considerati sono stati derivati i relativi indicatori di sintesi;

⁴ I valori I_k^N inoltre, se negativi, sono stati posti uguali a zero.

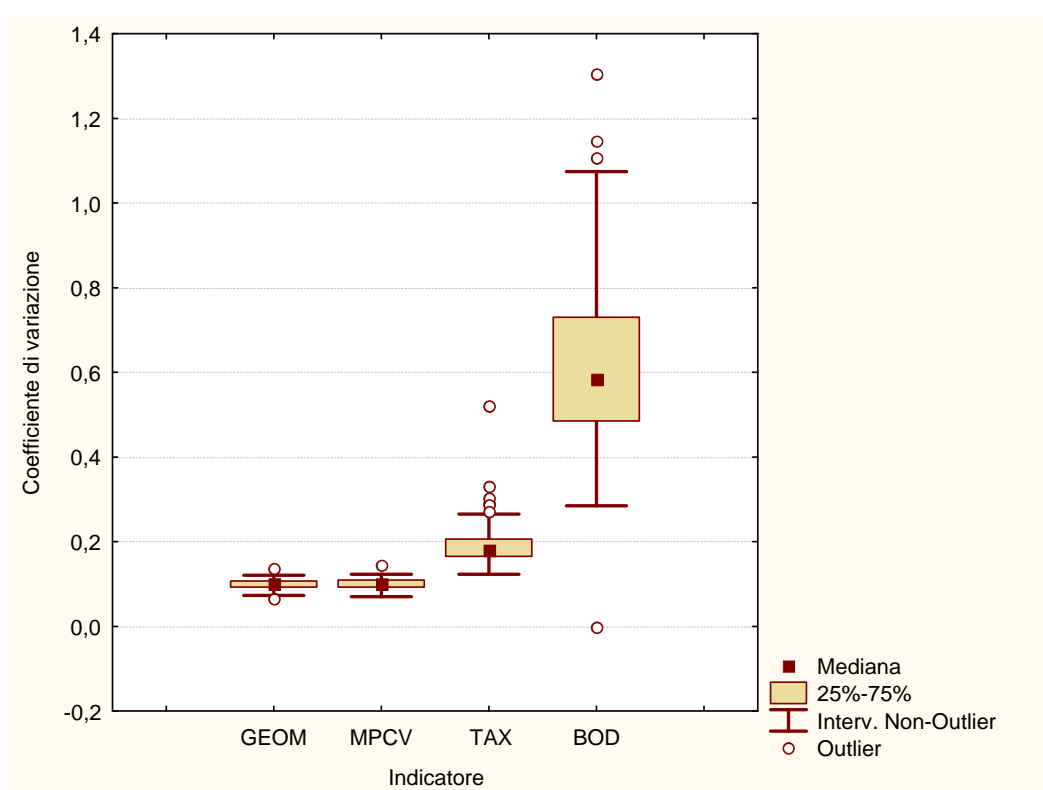
⁵ L'estrazione casuale è stata realizzata 100 volte.

- dato che i diversi indicatori, per loro stessa costruzione, non offrivano risultati confrontabili, presentando valori minimo e massimo differenti, si è calcolato, per ogni unità territoriale, il relativo coefficiente di variazione⁶ nelle diverse prove.

Dall'esame della Figura 2 si rileva come il metodo BOD presenti valori del coefficiente di variazione nettamente più alti rispetto agli altri metodi, specialmente al metodo MPCV ed alla media geometrica non ponderata (GEOM).

Inoltre, per il BOD questi valori risultano non correlati con il valore dell'indicatore di sintesi; per gli altri metodi, invece, si ottengono coefficienti di correlazione molto alti ($r \geq 0.8$) in quanto il disturbo, risultando proporzionale al valore dell'indicatore semplice, si riflette *direttamente* sul valore dell'indicatore composito.

Figura 2 Box-plot dei coefficienti di variazione degli indicatori di sintesi per i quattro metodi considerati ($\alpha=0,05$)



⁶ Il coefficiente di variazione rappresenta infatti, come è noto, un indice di dispersione che permette di confrontare misure di fenomeni riferite a unità di misura differenti, in quanto numero puro; esso viene definito come il rapporto tra la scostamento quadratico medio (σ) e il valore assoluto della media aritmetica (μ).

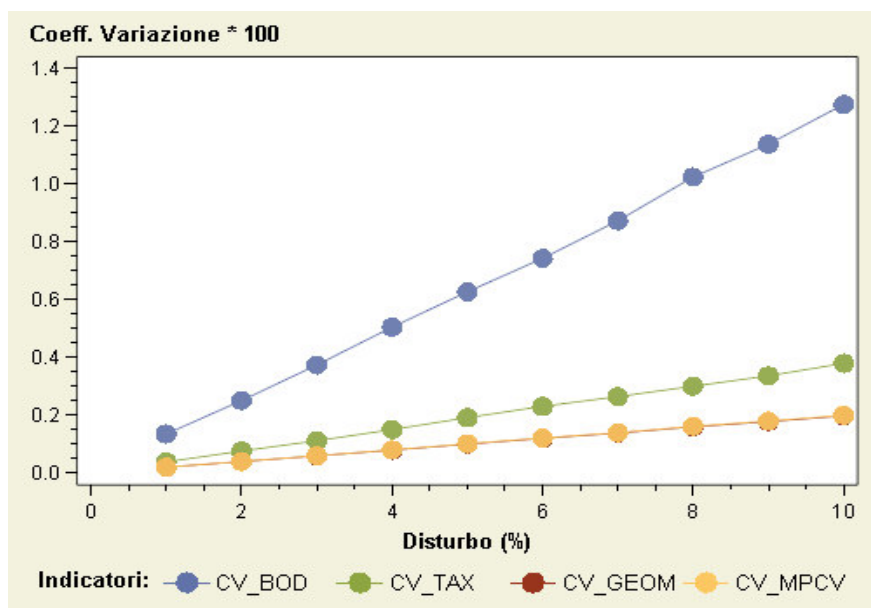
Ci si è, pertanto, chiesti se, e in quale misura, la variabilità crescesse al variare del disturbo aggiunto; l'analisi di robustezza è stata quindi ripetuta ipotizzando una variazione di α tra l'1% ed il 10%, e ricalcolando di volta in volta i relativi coefficienti medi di variazione.

Dalla Figura 3 si desume che la metodologia BOD è la più sensibile in termini assoluti rispetto agli altri metodi e che questa maggiore sensibilità sui valori dell'indicatore di sintesi cresce al crescere del disturbo aggiunto; questo perché la variabilità aggiunta sui dati si riflette indirettamente anche sui vincoli ai pesi nel problema di ottimizzazione lineare.

I metodi MPCV e GEOM reagiscono in modo molto simile (tanto da non rendere distinguibili, nella Figura 3, le due linee che li rappresentano).

L'aspetto più interessante dell'analisi, a questo punto, risiede nell'interrogativo se una maggiore variabilità degli indicatori di sintesi è in grado o meno di influire sugli ordinamenti dei relativi indici, ovvero se ad una maggiore variabilità dei dati corrisponda una maggiore media delle differenze assolute di rango rispetto ai relativi punteggi non disturbati.

Figura 3 Coefficiente di variazione medio degli indicatori di sintesi ottenuti con i quattro metodi al variare del disturbo casuale (600 estrazioni)



Facendo variare i dati relativi agli indicatori semplici ($\alpha = 5\%$) si possono ottenere, per ogni estrazione casuale e per ogni unità territoriale, le differenze assolute di ranking rispetto all'ordinamento ottenuto nel caso non disturbato e, dunque, per ogni unità, la media nelle diverse prove delle differenze assolute di ranking.

Nelle Figure 4 e 5 sono riportati, rispettivamente, i box-plot e gli istogrammi relativi alle distribuzioni di frequenze delle differenze medie assolute di rango per i differenti metodi di sintesi sperimentati.

Si rileva innanzitutto che tutti e quattro gli approcci qui considerati presentano una mobilità media di rango molto contenuta, segno evidente di una elevata robustezza dei risultati ottenuti fin dall'approccio più "semplice" (quello della media geometrica non ponderata).

Si nota inoltre che in questo caso è l'approccio BOD a presentare gli ordinamenti più robusti, ossia a risentire in misura minore dei disturbi simulati sui dati elementari.

Una motivazione di tale comportamento potrebbe risiedere nel fatto che il campo di variazione dell'indicatore, proprio per la procedura stessa della sua costruzione, risulta già nell'ordinamento originario (senza disturbi) più ampio che non negli altri metodi, il che consentirebbe di "assorbire" in qualche misura l'effetto della variabilità indotta dalle simulazioni effettuate.

Figura 4 Box-plot delle differenze medie assolute di rango per i quattro indicatori di sintesi

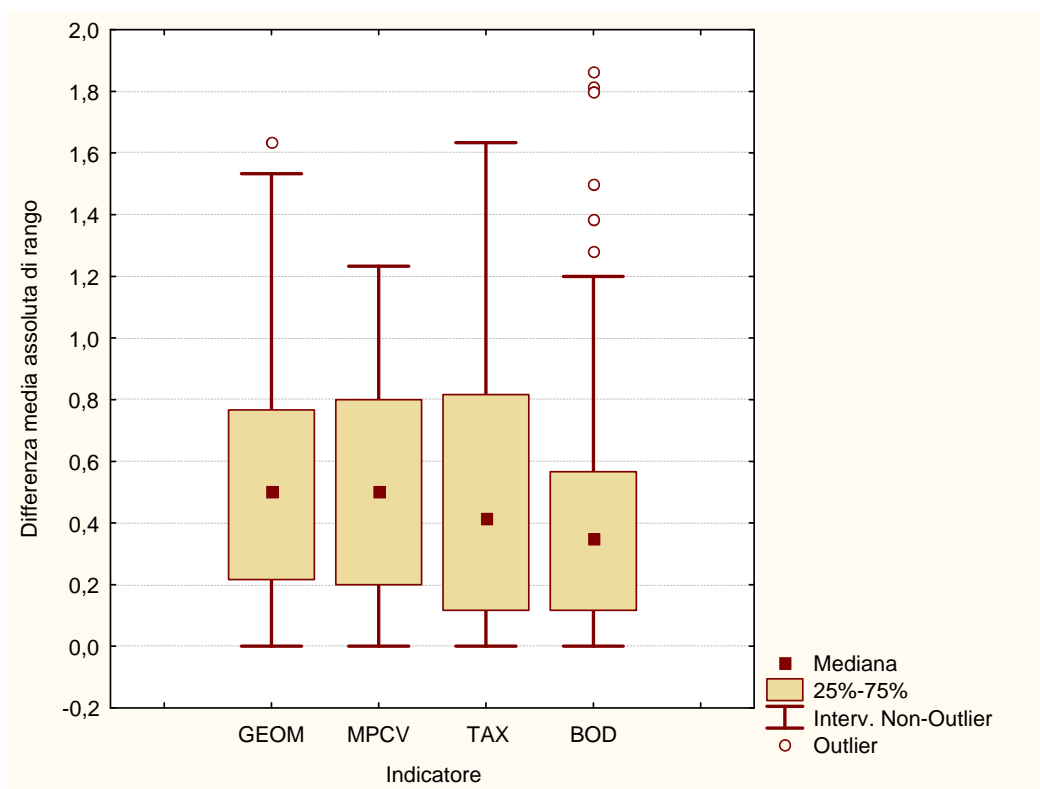
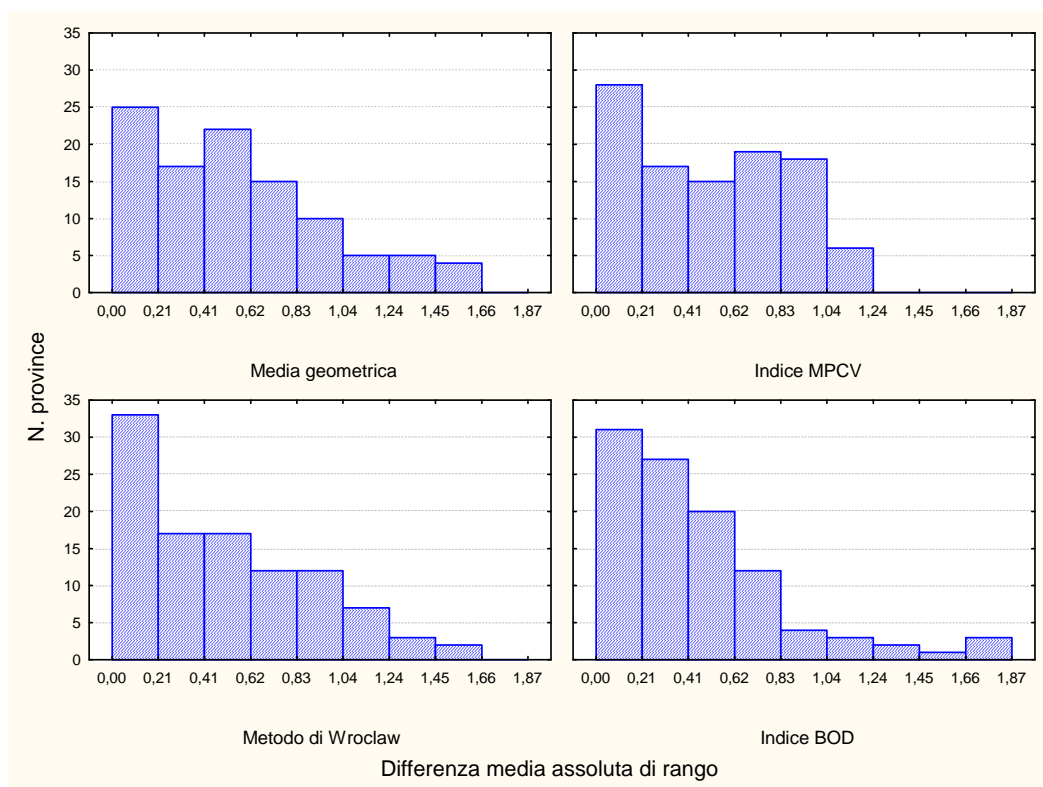


Figura 5 Distribuzione di frequenze delle differenze medie assolute di rango per i quattro indicatori di sintesi



4 COMMENTI CONCLUSIVI

Oggetto del lavoro presentato era il confronto tra diversi approcci di sintesi ponderata di indicatori elementari, confronto sorretto da una verifica di stabilità dei risultati ottenuti attraverso l'applicazione di tali approcci ad un set di 6 indicatori fisici di dotazione infrastrutturale articolati per le 103 province italiane. Va sottolineato che obiettivo del lavoro non era la scelta del metodo migliore – ritenendosi, al riguardo, che ogni metodo possa essere considerato migliore rispetto ad uno specifico obiettivo dell'analisi – ma, più semplicemente, la verifica del grado di affidabilità “statistica” dei risultati ottenuti con i diversi approcci.

Al termine del lavoro svolto, si possono formulare i seguenti commenti conclusivi:

- la limitazione dell'applicazione ad un set ristretto di infrastrutture (quelle stradali e quelle ferroviarie) non sembrerebbe aver inficiato la correttezza del confronto effettuato. Gli indicatori elementari sottoposti alle diverse procedure di sintesi presentano infatti un rilevante grado di variabilità, quale verosimilmente avrebbe caratterizzato un set di categorie infrastrutturali meno omogenee di quelle qui considerate;

- ai dati elementari è stata dapprima applicata una procedura di *sintesi non ponderata*, assunta quale approccio-base per i successivi confronti. Questi ultimi sono stati effettuati sulla base di tre distinti approcci di *sintesi ponderata*, scelti tra quelli più consolidati in letteratura (l'approccio del metodo tassonomico) o tra quelli applicati in recenti lavori specificamente dedicati al tema della dotazione infrastrutturale (metodo delle penalità per coefficiente di variazione e metodo del *benefit of the doubt*);
- una prima verifica descrittiva mostra un'elevata concordanza tra le graduatorie ottenute dall'applicazione dei metodi di sintesi considerati: indice di cograduazione sempre maggiore di 0,95; spostamenti in graduatoria tra l'una e l'altra applicazione in media molto contenuti, compresi tra 1 e 6 posizioni, su 103 unità territoriali considerate;
- la verifica è stata successivamente approfondita attraverso un'analisi di sensitività (in termini di reazione al variare del set originario di dati) e un'analisi di robustezza (in termini di reazione a seguito dell'intervento di disturbi casuali sui dati originari), entrambe le analisi volte a saggiare la stabilità "statistica" dei risultati ottenuti con i diversi approcci di ponderazione;
- pur nella differenza tra i risultati di dettaglio di ciascuna delle due analisi, le verifiche effettuate – valutate attraverso le differenze di rango tra le graduatorie ottenute dagli approcci considerati – conducono ad un giudizio di forte convergenza dei risultati ottenuti. In altri termini, la geografia infrastrutturale che si presenta come esito finale dei diversi approcci risulta statisticamente robusta, evidenziando ancora una volta una configurazione della dotazione infrastrutturale (qui dei soli dei trasporti terrestri) mediamente alquanto sbilanciata a sfavore di gran parte del territorio meridionale, nonché di alcune province prevalentemente appartenenti ai confini della ripartizione nord-orientale.

Riconoscimenti

Il lavoro è stato condotto in stretta collaborazione tra gli Autori. Quanto alla redazione del testo, Claudio Mazziotta ha scritto i parr. 1 e 4; Matteo Mazziotta i parr. 2.1, 2.2, 3.1 e 3.2; Adriano Pareto i parr. 2.3 e 3.3; Francesco Vidoli i parr. 2.4 e 3.4.

Riferimenti bibliografici

- Aiello F., Attanasio M. (2004), How to transform a batch of simple indicators to make up a unique one?, Società Italiana di Statistica, Atti della XLII Riunione Scientifica, CLEUP, Padova.
- Aureli Cutillo E. (1996), *Lezioni di statistica sociale. Parte seconda, sintesi e graduatorie*, C.I.S.U., Roma.
- Biehl D. (editor) (1986), *The Contribution of Infrastructure to Regional Development*, Commission of the European Communities, Infrastructure Study Group, 2 vols., Bruxelles.
- Brunini C., Paradisi F., Terzi S. (2004), Banca dati delle infrastrutture: una proposta di metodo per la sintesi di indicatori, ISTAT, Progetto interdipartimentale "Informazione statistica territoriale e settoriale per le politiche strutturali 2001-2008", mimeo, Roma.
- Cherchye, L., Moesen, W., Van Puyenbroeck, T., (2004), "Legitimately Diverse, yet Comparable: on Synthesizing Social Inclusion Performance in the EU", *Journal of Common Market Studies*, 42.
- Cherchye, L., W. Moesen, C.A.K. Lovell e T. Van Puyenbroeck (2005), "One Market, One Number: A Composite Indicator Assessment of EU Internal Market Dynamics", *CES DP Series 05.16*, KU Leuven.
- Cherchye, L., W. Moesen, Rogge N., Saisana M., Saltelli A., Liska R., Tarantola S. (2006), "Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index", *Public Economics Working Paper Series*, n° ces0613, Katholieke Universiteit Leuven, Centrum voor Economische Studiën.
- Di Palma M. e Mazziotta C. (2003), "Infrastrutture, competitività e sviluppo: il caso italiano", *Economia italiana*, n. 1.
- Istituto Nazionale di Statistica (2006), *Le infrastrutture in Italia. Un'analisi provinciale della dotazione e della funzionalità*, Roma.
- Mazziotta C. (1998), "Definizione di aree e indicatori per la misurazione della dotazione di infrastrutture", in Istituto Guglielmo Tagliacarne, *Statistica e territorio. Esperienze e nuovi percorsi di ricerca per l'analisi delle economie locali*, F. Angeli, Milano.
- Mazziotta C. (2006), "Infrastructure, Services and Competitiveness. The Italian Case", *Atti della XLIII Riunione Scientifica della Società Italiana di Statistica*, Torino.
- Mazziotta C., Vidoli F. (2007), "La costruzione di un indicatore sintetico ponderato: un'applicazione della procedura *Benefit of the Doubt* al caso della dotazione infrastrutturale in Italia", in *Lo sviluppo regionale nell'Unione Europea - Obiettivi, strategie, politiche*, Atti della XXVIII Conferenza Italiana di Scienze Regionali, Bolzano.
- Mazziotta M., Pareto A. (2007), *Un indicatore sintetico di dotazione infrastrutturale: il metodo delle penalità per coefficiente di variazione*, in *Lo sviluppo regionale nell'Unione*

Europea - Obiettivi, strategie, politiche, Atti della XXVIII Conferenza Italiana di Scienze Regionali, Bolzano.

Melyn W., Moesen H. P. (1991), "Towards a synthetic indicator of macroeconomic performance: unequal weighting when limited information is available", KU Leuven Center for Economic Studies, *Public Economics Research Paper 17*, Leuven.

Nardo M., Saisana M., Saltelli A., Hoffman A. e Giovannini E. (2005), "Handbook on constructing composite indicators: methodology and user guide", *OECD Statistics Working Papers JT00188147*, Paris.

Quirino P. (1990), *Indicatori socio-culturali a livello regionale*, Cresa, Collana di studi economici, L'Aquila.

Schifini D'Andrea S. (1982), «Indagini sul livello di sviluppo regionale in Italia (un'applicazione comparativa)», in AA. VV., *Le statistiche dello sviluppo*, Edizioni Scientifiche italiane, Napoli.

APPENDICE

Prospetto 1 Indicatori di dotazione delle infrastrutture dei trasporti stradali e ferroviari

Territorio	ITR012_1999	ITR013_2000	ITR014_1996	ITR015_2003	ITR064_2004	ITR065_2004
Agrigento	19,38	31,73	21,07	0,00	0,00	26,27
Alessandria	23,47	49,81	11,04	51,93	43,27	46,65
Ancona	22,36	43,48	13,97	29,64	46,23	25,02
Aosta	8,62	15,30	8,00	33,71	0,00	0,00
Arezzo	18,61	30,26	14,26	16,84	47,91	0,00
Ascoli Piceno	27,68	80,41	12,60	23,57	22,31	0,00
Asti	37,38	77,71	11,25	25,88	35,57	24,08
Avellino	33,63	36,06	25,33	32,63	0,00	11,34
Bari	24,58	34,38	15,18	22,44	25,92	10,43
Belluno	13,95	10,51	19,85	3,97	0,00	0,00
Benevento	34,07	52,06	15,55	0,00	9,24	26,07
Bergamo	25,17	39,37	10,58	11,79	8,31	28,76
Biella	25,09	63,67	14,57	0,00	0,00	0,00
Bologna	20,61	29,17	12,02	39,24	58,26	21,75
Bolzano-Bozen	8,37	16,27	11,07	15,77	17,22	12,93
Brescia	23,46	22,07	13,29	20,36	15,67	9,18
Brindisi	27,19	41,04	13,32	0,00	36,14	29,17
Cagliari	19,18	18,85	13,52	0,00	2,36	0,00
Caltanissetta	11,57	53,39	12,59	10,05	0,00	35,39
Campobasso	24,19	40,17	26,99	12,41	7,92	8,93
Caserta	29,03	49,75	17,73	32,02	72,59	16,95
Catania	18,23	37,98	13,34	20,78	10,93	12,76
Catanzaro	31,19	53,09	22,58	19,86	16,08	1,61
Chieti	36,98	53,32	23,64	29,75	15,23	23,65
Como	32,89	38,04	10,56	22,59	21,51	0,00
Cosenza	26,73	33,69	23,22	20,78	15,73	20,53
Cremona	17,01	41,29	14,57	10,67	0,00	68,44
Crotone	16,57	36,70	22,54	0,00	0,00	0,00
Cuneo	17,11	40,78	11,36	11,85	10,60	20,84
Enna	16,23	30,63	16,24	21,74	0,00	29,39
Ferrara	18,63	27,14	8,21	10,49	6,24	15,44
Firenze	19,10	27,01	15,00	27,43	48,88	1,45
Foggia	16,18	32,35	13,08	23,79	14,50	8,36
Forlì-Cesena	19,94	38,58	14,89	18,09	15,36	3,34
Frosinone	32,40	49,70	16,15	25,96	26,23	0,00
Genova	32,33	43,37	23,55	80,56	95,37	14,90
Gorizia	19,36	27,50	22,75	33,90	78,48	4,53
Grosseto	12,18	32,36	13,10	0,00	20,40	0,00
Imperia	38,76	65,43	13,76	53,81	26,82	16,47
Isernia	23,54	45,83	19,17	0,00	0,00	19,01
L'Aquila	20,46	24,96	18,25	25,98	0,00	23,22
La Spezia	35,47	56,45	24,83	68,79	71,11	21,12
Latina	18,62	42,13	13,91	0,00	39,44	12,58
Lecce	37,91	69,89	18,23	0,00	1,29	7,51
Lecco	29,81	46,62	10,66	0,00	15,65	73,25
Livorno	19,07	41,50	15,52	35,80	84,45	10,93
Lodi	15,91	50,87	15,98	50,50	61,58	9,86
Lucca	39,50	28,93	12,47	42,53	10,77	27,10

Territorio	ITR012_1999	ITR013_2000	ITR014_1996	ITR015_2003	ITR064_2004	ITR065_2004
Macerata	18,42	47,69	14,13	6,49	5,37	0,00
Mantova	18,86	35,92	12,91	16,12	0,00	42,12
Massa-Carrara	36,88	46,99	13,40	49,29	27,52	25,12
Matera	16,09	30,40	18,28	0,00	0,00	33,60
Messina	30,10	81,84	15,12	48,56	15,50	44,87
Milano	52,05	20,84	15,89	93,98	142,08	26,06
Modena	27,23	27,15	17,07	19,08	12,10	19,27
Napoli	47,69	110,51	25,10	87,61	96,83	30,46
Novara	27,59	46,65	15,09	76,00	44,01	63,35
Nuoro	16,29	17,57	13,49	0,00	0,00	0,00
Oristano	23,06	34,21	9,58	0,00	0,00	0,00
Padova	28,64	51,36	12,05	34,55	59,37	23,66
Palermo	19,47	44,11	15,46	29,57	10,49	30,73
Parma	23,64	30,13	13,48	27,05	13,76	28,66
Pavia	17,53	54,98	12,68	31,67	36,09	8,76
Perugia	26,36	31,83	16,53	1,56	8,29	24,45
Pesaro e Urbino	26,85	47,78	11,10	15,04	13,76	0,00
Pescara	33,23	50,37	20,41	48,52	5,87	38,58
Piacenza	20,92	33,12	13,63	34,60	19,51	17,81
Pisa	19,84	31,37	14,50	10,18	33,33	26,34
Pistoia	32,47	36,49	18,55	29,85	17,81	67,27
Pordenone	17,56	28,64	9,06	14,06	12,28	0,00
Potenza	22,10	27,62	19,31	4,47	3,74	9,62
Prato	52,76	14,62	15,06	26,28	56,14	0,00
Ragusa	18,16	57,12	9,29	0,00	0,00	0,00
Ravenna	20,74	36,09	15,28	25,77	12,34	72,70
Reggio di Calabria	32,08	45,71	20,67	23,00	22,17	0,22
Reggio nell'Emilia	21,23	40,07	12,60	17,66	12,78	3,20
Rieti	23,54	41,57	16,62	11,89	11,86	0,00
Rimini	45,95	53,11	14,80	55,13	52,60	24,92
Roma	27,33	34,85	15,62	61,42	71,75	31,38
Rovigo	14,25	45,73	9,27	13,63	16,13	0,00
Salerno	24,48	43,29	18,93	33,25	38,17	14,61
Sassari	12,88	26,65	10,38	0,00	0,00	0,00
Savona	36,32	37,86	17,93	67,58	31,54	62,28
Siena	10,42	30,43	16,41	8,77	8,88	0,00
Siracusa	15,07	60,94	11,43	4,55	3,12	28,98
Sondrio	13,33	9,89	9,87	0,00	0,00	28,86
Taranto	25,85	61,56	13,75	10,95	5,39	36,05
Teramo	29,54	79,66	19,94	44,50	25,49	12,87
Terni	19,43	34,03	17,67	23,04	60,31	8,26
Torino	22,80	38,80	12,23	40,75	29,19	10,98
Trapani	19,29	38,03	12,96	50,82	0,00	0,00
Trento	17,38	23,60	13,74	11,34	11,51	0,00
Treviso	27,98	52,93	12,03	30,65	33,19	19,70
Trieste	56,75	63,61	41,54	28,80	238,29	120,98
Udine	15,66	25,73	17,05	31,83	37,22	11,26
Varese	42,79	50,31	19,94	34,37	39,31	69,34
Venezia	19,28	36,13	12,46	40,10	46,24	14,06
Verbano-Cusio-Ossola	12,21	20,55	15,03	0,00	21,01	21,57
Vercelli	14,77	35,46	14,56	49,71	26,34	13,39
Verona	28,44	49,87	10,00	36,72	45,81	17,37
Vibo Valentia	38,58	65,82	22,55	31,59	39,75	55,69
Vicenza	30,16	35,95	14,55	26,53	20,44	5,10
Viterbo	18,18	33,67	9,99	6,29	33,86	18,80
Media aritmetica	24,76	41,19	15,56	24,88	27,03	19,83
S.q.m.	9,64	16,29	4,89	20,86	33,12	20,88
C.V.	0,39	0,40	0,31	0,84	1,23	1,05

ABSTRACT

The synthesis of different indicators expressed in only one representative measure of a set of phenomena is a subject known in literature, especially in the social statistics. Frequently the synthesis is calculated without weighting, or rather it is assumed, implicitly, that the phenomena considered have the same importance and they can be considered on the same level. This assumption reduces the risk of the arbitrariness but it could lead biased, or not completely reliable, results.

However the issue is more complicated when there are two dimensions to summarize: the sectorial dimension (in this paper, the different infrastructural categories) and the territorial dimension (the Italian provinces). The problem – in particular, computing composite indicators of infrastructural endowment – is dealt with different approaches based on the sectorial and territorial variability in order to determine the weights for a set of elementary indicators.

The aim of this paper is to compare some approaches, to verify the convergence and the robustness (reliability) of the results obtained.